Analysis of data from the Atmospheric Visibility Monitoring (AVM) program

Muthu Jeganathan[a] and Neema Jalali[b]

[a]Jet Propulsion Laboratory, MS 161-135, 4800 Oak Grove Dr., Pasadena, CA 91109

[b]California Institute of Technology, Pasadena, CA 91125

ABSTRACT

The Atmospheric Visibility Monitoring (AVM) program at the Jet Propulsion Laboratory has been in place for the last few years to obtain atmospheric transmission statistics data to support free-space optical communications experiments and missions. Atmospheric transmission data is collected through a set of three autonomous systems, all located in the south-western U.S., that observe selected stars throughout the year. Data from these three sites are collected and processed on a regular basis to obtain cumulative distribution functions of atmospheric attenuation for different spectral windows of interest. In this paper, we describe recent work on creation of a database using Microsoft Access to analyze atmospheric transmission data collected by the AVM project. The database, which replaces time-consuming Matlab programs, offers a rapid and extremely easy way to extract and analyze AVM data.

Keywords: Atmospheric visibility, Atmospheric transmission, Atmospheric attenuation, Cloud-cover statistics, Free-space Optical Communication, Laser Communication, Link Budget, AVM, CDF, PDF

1. INTRODUCTION

An Earth-to-space optical communication link requires a good understanding of the atmosphere as the laser beam has to propagate through it. The atmosphere not only attenuates the light wave but also distorts and bends it. Attenuation is primarily the result of absorption and scattering by molecules and particulates suspended in the atmosphere. Distortion, on the other hand, is caused by atmospheric turbulence through index of refraction fluctuations. Attenuation affects the mean value of the received signal in an optical link where as distortion results in variation of the signal around the mean. Often, atmospheric attenuation, or in the extreme case cloud cover, can be the limiting factor in an optical communication link through the atmosphere. In this article we shall, therefore, concentrate on the statistics of atmospheric attenuation and ignore refraction as well as the physical mechanisms that are responsible for radiation losses in the atmosphere.

The atmospheric visibility monitoring (AVM) project was initiated by the Optical Communications Group (OCG) of the Jet Propulsion Laboratory to determine experimentally the statistics of atmospheric transmission at different optical frequencies. As part of the project three AVM units were built and deployed around the south-western U.S. Using calibrated telescopes, these AVM systems measure intensities of known stellar objects to determine atmospheric attenuation statistics. Post-processing of the collected data provides atmospheric visibility statistics. Such data is necessary to figure out, for example, the probability with which data can be relayed back to Earth from satellites or spacecrafts using laser communications. The model is expected to be used in link margin analyses for optical communications channels. Atmospheric transmission data from multiple sites will be helpful in determining the probability that at least one of several sites can see a spacecraft's downlink. The RF community has long gathered similar data in the RF spectrum. For example, Ka-band beacons on board the ACTS satellite are monitored by several ground stations to determine the effects of rain and clouds on propagation of radiation around 20 GHz [1].

In this article we describe recent development and construction of a database to maintain and analyze AVM data. We begin with a brief description of the AVM stations and their capabilities. Details of the contents of the data files produced by each AVM station leads into the discussion of the database. A section is devoted to present results from the database showing its power and usefulness. We conclude with possible improvements to the database and future plans.

Send correspondence to Muthu Jeganathan at muthu@altair.jpl.nasa.gov

2. SYSTEM DESCRIPTION

As mentioned earlier, there are now three AVM systems in operation, all in the south-western United States. These are at 1) Table Mountain Facility near Wrightwood, California; 2) Mt. Lemmon near Tucson, Arizona; and 3) Goldstone near Barstow, California. Each AVM unit is a self-contained (in a 8 ft long x 8 ft wide x 8 ft high) an enclosure) autonomous system that continuously observes selected stars and collects data. At the heart of each AVM unit is a 386-based computer which controls telescope motion, camera, filter wheel, etc. Stars are observed through a 10-inch f10 telescope with one of six filters (three narrow band filters at 532 nm, 860 nm & 1060 nm and three broadband astronomical filters V, R & I). Since the current CCD is not sensitive at 1060 nm, data with that filter is not usable. Star intensities are recorded using a cooled, slow scan CCD camera. In addition to star intensities, the AVM units record weather information such as temperature, wind speed and humidity.

The AVM stations are programmed to observe one star (out of three) for 15 minutes. The three stars are selected so that they are separated by approximately 8 hours in RA. That is, each AVM telescope observes the stars as they rise and tracks them until they set. Normally only one of the three stars is up at any given time. This ensures that nearly all the time, the same star is observed from all three sites, giving the needed site-diversity data.

During each 15 minutes, six images of the stars are taken with the six different filters. Each image is stored with weather and status information in a FITS format file. At any given site, about 550 observations are made each day. A computer at JPL calls the AVM PC over standard phone line and picks up the data collected over the last 24 hours. A typical day's data for each site is about 1.4 Mbytes. At the end of each month, the data collected for that month is stored on a CD using a recordable CD player. Archival of data on CDs allows easy access to the data from any computer. The analysis of the data is the focus of this paper. In-depth information on AVM hardware and operational details can be found in References [2,3]. Over the past year, we have also undertaken an effort to upgrade the computer, CCD camera and software to increase system reliability [3].

Previous data processing scheme, using Matlab, was adequate to analyze quarterly data and produce cumulative distribution functions (CDFs) of atmospheric attenuation for each quarter. The method, however, was not easily extended to process several years of data or to extract information not originally intended from the gathered data. As stated earlier, each AVM unit makes 24 observations per hour. This amounts to more than 200,000 observations from each site per year or well over half-a-million observations per year from all three sites. After struggling with the Matlab code, we decided to use a commercial database program to do the data processing and analyses. Description of the database and results from it follows.

3. AVM DATABASE

As mentioned earlier, at the end of each month, all AVM data is archived onto a CD-ROM. At this point, the data is still in its raw form and needs to be "reduced" before entering them into a database. Although many data fields are readily usable (see Table 1 for a sample observation file), there are many fields that require further processing. The star image, for instance, is stored using a FITS data format and must be processed before any useful data can be extracted. This process, commonly referred to as reduction, is achieved using the AVM Data Reduction Program.

The program begins by reading each individual observation file from the CD-ROM and processing the star image contained within. More specifically, the program correlates the image against an expected point spread function (PSF). It is assumed that the PSF, or the long-exposure intensity distribution at the focal plane, is a Gaussian because of atmospheric turbulence. The correlation process yields the location of the star inside the image. The program then calculates the background noise of the image by averaging the pixel values around the border. This amount is subtracted from *all* the pixels so that the contribution from dark or background noise is reduced. By summing the pixel values in the neighborhood of the star's center, we finally obtain the desired information, i.e. a number E that is proportional to the total collected star energy.

Though the calculations for E are straightforward, the process does not always result in a meaningful answer because of significantly attenuated star light or background solar radiance. A set of heuristic algorithms are used to determine the probability that a star is actually present in the image. These algorithms generate a flag which represents the confidence in the calculations. A flag of 4 indicates high confidence while a value of zero means no star is in the image. The fact that intermediate flag numbers (especially 2 and 3) do not often occur indicates validity of the approach. When no image is present in the observation file because of unfavourable weather conditions, the flag is set to -1. Whenever the flag is 2 or lower, the atmospheric transmission is set to zero.

Table 1. A sample AVM observation file that is created after every observation. At the end of this text file would be the CCD image. Compressed version of these raw files are archived on CDs.

```
SIMPLE  =                    T /
BITPIX  =                   16 /
NAXIS   =                    2 /
NAXIS1  =                   42 /
NAXIS2  =                   42 /
COMPRESS= 'SPECTRASOURCE'      /
TELESCOP= 'TMF     '           /
STAR_NUM=                 8781 /
FILTER  =                    1 /
UT_TIME = ' 14:24:49.97'       /
EXP_TIME=                41840 /ms
MAX_PIX =                 3566 /
MAX_XLOC=                   78 /
MAX_YLOC=                    4 /
MIN_PIX =                    0 /
STAR_RA = ' 07:45:06.1'        /hours, minutes, seconds
STAR_DEC= '  28:02:05'         /degrees, minutes, seconds
APP_RA  = ' 23:04:37.9'        /hours, minutes, seconds
APP_DEC = '  15:11:39'         /degrees, minutes, seconds
CEN_RA  = ' 23:04:38.8'        /hours, minutes, seconds
CEN_DEC = '  15:11:40'         /degrees, minutes, seconds
IMG_TYPE= 'SUBTRACTED'         /
DATE_OBS=       2.45029550e+06 /
ZEN_ANG =          5.8416e+01 /degrees
AZIMUTH =         -1.8746e+00 /degrees
OUT_TEMP=          2.4600e+01 /degrees C
CART_TMP=          1.5800e+01 /degrees C
REL_HUM =                   18 /
WIND_DIR=                  214 /degrees
WIND_SPD=                    0 /m/s
CCD_TEMP=         -8.2000e+00 /degrees C
CCD_C_ST= ' ON     '           /(on/off)
UPS_STAT= '   OK '             /(ok/alarm/bat)
ROOFSTAT= ' Open '             /(open/close/fail)
PRECIPST= ' NO '               /(yes/no)
END
```

Table 2. Fields in the AVM database. All but the last of the 38 fields are either directly or indirectly from the observation file shown above.

Info	Fields	Notes
site	name	Site name
time	year, month, day, hour, minutes, seconds, hundredths of seconds	Observation time
star	Bright Star (BS) number	
filter	number	Corresponds to type of filter
weather	temperature in Celsius, relative humidity in %, wind speed in m/s, wind direction, rain	Weather information
data	exposure time in seconds, zenith angle in degrees, azimuth angle in degrees, collected energy, average background pixel value, Q1, Q2, Q3, Q4, flag	Q1-Q4 and flag indicate confidence in energy calculations
image	row number, column number, pixel value	Information on brightest pixel
status	cart temp., CCD temp., CCD cooler status, UPS status, roof status, apparent RA, apparent DEC, change in RA, change in DEC	System health monitors
misc.	mm-dd-yyyy	Date data was reduced

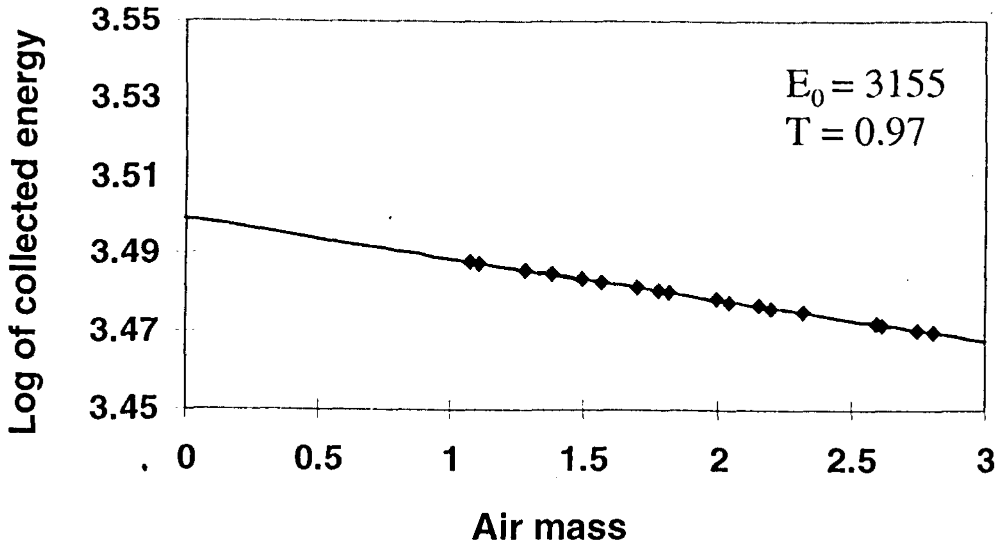

Figure 1. Calibration plot showing log of measured intensity vs. air-mass. Markers show data points while the solid line represents the fit. The calibration is for Mt. Lemmon AVM system for the 860 nm filter and star #8781.

The overall task of the AVM Data Reduction Program is to convert each observation from an entire file to a single line of ASCII text. Each of these lines of text contains 38 tab-separated data fields, including star energy and flag number as calculated in the process described above. Other fields are either transferred directly from the observation file or converted via simple arithmetic means. Table 2 lists the various fields that are used in the database. This reduction process is useful because it not only reduces the overall size of the data but also places it in a format suitable for entering into the AVM Database Program.

After reducing several months of AVM data, it is loaded into the Microsoft Access database. We chose Access over other database programs, such as the ones from Oracle Corp., primarily because Access comes standard with the Professional version of the popular Microsoft Office suite. Furthermore, Access is lot easier to use than the more capable Oracle software. So far, we have been quite pleased with the capabilities of Access. The database contains the 38 data fields discussed above, plus two new ones - one which assigns each record a unique identification number (ID), and the other which tells whether that particular record was taken during daytime or nighttime. The latter had to be done because the AVM system, in its current configuration, can not distinguish stars from the bright solar background during daytime. That is, we only use data collected at night for the visibility CDFs. The day/night calculation is based on the definition of civil twilight and is calculated using the site location, date and time.

Before we can use the database, however, we need to know a set of calibration values to determine the atmospheric transmission from the measurements made by the CCD. The calibration values essentially provide what an AVM system would have observed if it were outside the Earth's atmosphere (i.e. in the absence of any atmosphere). Since the throughput of the AVM optical system and the CCD response were not adequately characterized in-lab, the calibration values must be extracted from the data itself. This is easily done by realizing that the atmospheric transmission and the angle of observation obey a simple relationship. Let I_0 be the intensity of light from a star incident above the Earth's atmosphere, θ be the zenith-angle of observation and η be the transmission at zenith (i.e. looking straight up) at a particular site. Then the light intensity, I, from a star, assuming a plane-parallel atmosphere, is [3]:

$$I = I_0 \eta^m \tag{1}$$

The factor m is often called the air-mass and is approximately equal to $\sec\theta$. The air mass is a measure of the "amount" of atmosphere that light travels through. For $\theta = 0$, m is 1 and for $\theta = 60$ degrees, m is 2. For zenith-angles around 70 degrees or above, the $m = \sec\theta$ expression for air-mass fails because the Earth curves around and the plane-parallel atmosphere assumption is no longer valid. Thus, we use the following formula for calculating the air mass [4]:

$$m = \sec\theta - 0.0018167(\sec\theta - 1) - 0.002875(\sec\theta - 1)^2 - 0.0008083(\sec\theta - 1)^3 \tag{2}$$

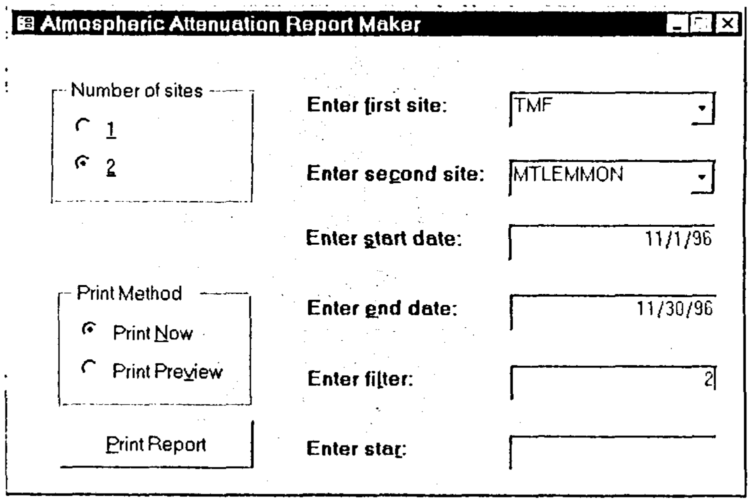

Figure 2. Query dialog box for obtaining single- or two-site CDFs

which extends the accuracy of air-mass calculations for θ up to about 75 degrees.

Through a change of variables ($y = \ln I$, $x = m$ and $y_o = \ln I_0$), the equation for the observed intensity can be written as a linear equation:

$$y = y_0 + ln(\eta)x \tag{3}$$

The observed data can then be fit to obtain y_0 and $\ln(\eta)$ from which calibration values, I_0, can be determined. Since the radiation spectra of stars vary and the AVM systems are not identical, calibration values must be obtained for each filter, star and site. It must be noted that the above expressions assume that the atmosphere (or η) is constant as a star is observed at different angles. This is, of course, not the case. Calibration values can still be determined by realizing that over a period of several months, a given star will be observed at several angles through *clear* atmosphere (i.e. on photometric nights). In other words, for a given air-mass we choose a data point with the largest measured value of star-light. This process will yield the calibration value as well as the maximum value of η for a particular filter, star and site.

The actual calibration in Access is done as follows. We first divide the interval of air-mass between 1 and 4 into smaller bins, typically 0.1 air-mass in size. Using the features of Access, we then find the highest measured value of star light in each bin. The resulting group of data points is expected to lie in a straight line as in Eq. 3. A robust line fitting algorithm based on minimum absolute deviation (see Section 14.6 of [5]) is used instead of the common χ-squared or least squares algorithm because of the possibility that in some bins a star may not have been observed under the best atmospheric conditions. The relevant routines in Ref. [5] were converted to Visual Basic for use in Access. From the fit, we extrapolate what the system would have observed at zero air-mass (or above the atmosphere) to give the desired calibration numbers (see Figure 3). In addition to the AVM database, we thus have a table of calibration values for each filter, star and site.

4. RESULTS

Once the reduced data is loaded and the calibration table is set up in Access, the database is ready for use. The power of Access comes from the fact that one can easily and rapidly extract data by imposing certain bounds on one or more of the various fields. In database language, this process of selectively extracting data is called a query. In the process of using the database, numerous queries have been defined to access the most common information. The most important query is the one which generates a cumulative distribution function (CDF) of atmospheric attenuation given a location, filter, and a time period. In Access, the query can be linked to a dialog box using Visual Basic for Applications (VBA) that makes obtaining CDFs trivial. Figure 2 shows such a dialog box.

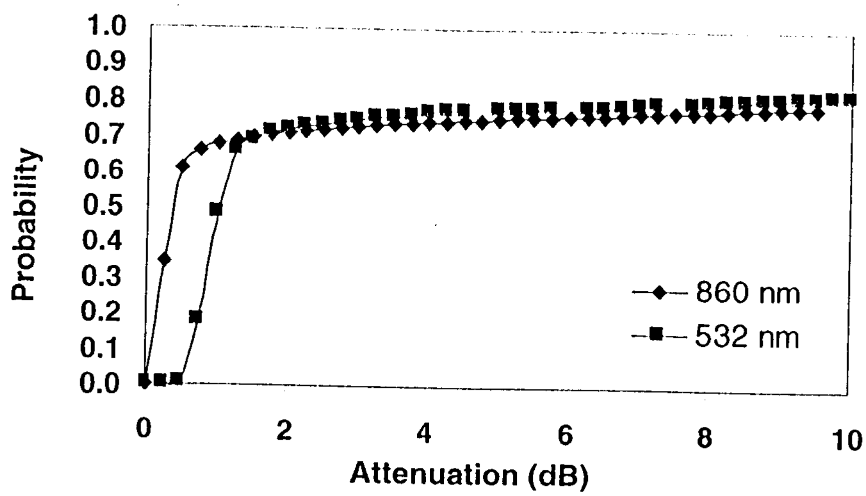

Figure 3. Attenuation CDF for Mt. Lemmon during November 1996 at 532 and 860 nm.

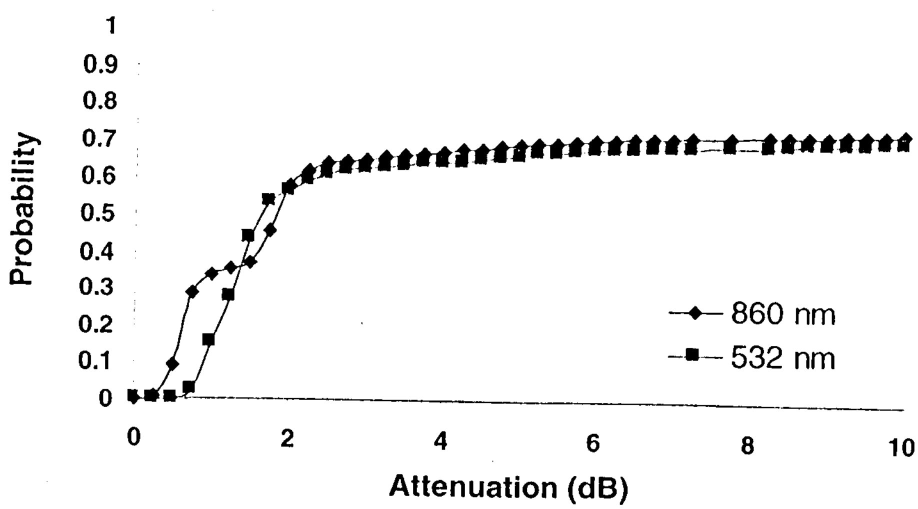

Figure 4. Visibility CDF for TMF during November 1996 at 532 and 860 nm.

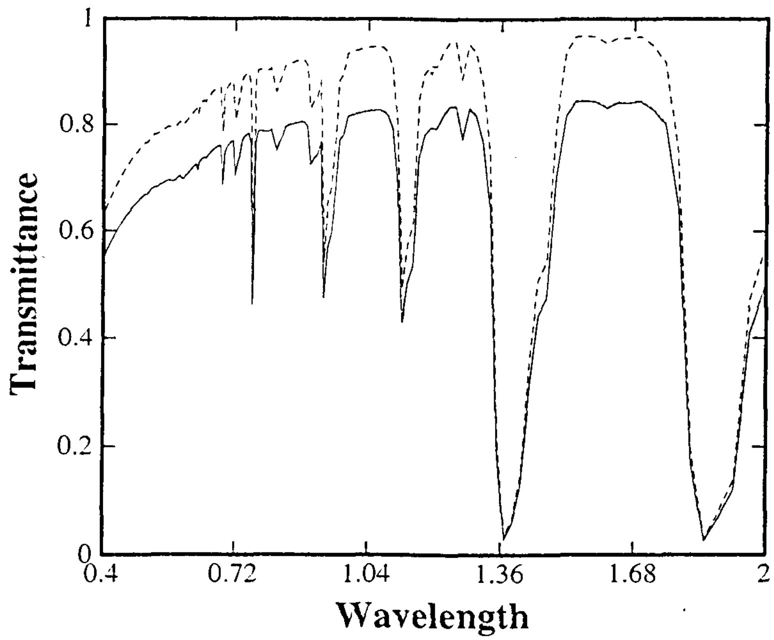

Figure 5. Atmospheric transmission at zenith as a function of wavelength for a 2 km altitude site generated from MODTRAN [6]. The two curves are for clear sky and high cirrus clouds. Note: both TMF and Mt. Lemmon are at altitudes slightly higher than 2 km.

The queries use the following equation to determine the transmission, T, using values from a particular observation:

$$T = \left(\frac{E}{E_0}\right)^{1/m}$$

where E is the collected energy, E_0 is the calibration value, and m is the air mass. The energy values are adjusted to have a constant exposure time. Note that the transmission values are normalized to zenith through the $1/m$ power. Indeed all CDF plots that we generate are for zenith attenuation. It is convenient to define attenuation, A, as follows:

$$A = 10\ \log(T)$$

An attenuation of 3 dB ($A = -3$ dB) thus refers to a transmission of 50%. Figure 3 and Figure 4 show the attenuation CDF for TMF and Mt. Lemmon, respectively. The point where the CDF intersects the horizontal axis indicates the highest transmission observed at the site. These values should be compared with the atmospheric transmission curves shown in Figure 5. The unusual step in the CDF for TMF at 860 nm is being investigated

A multi-site joint distribution can also be produced from the database. For joint CDFs, we use the highest atmospheric transmission from two or more sites in a given 15-minute time interval. Figure 6 shows the joint CDF for TMF and Mt. Lemmon for the 532 nm filter. Note that over 80% of the time at least one of the two sites has attenuation less than 2 dB. Joint CDFs are extremely useful in communications link to determine the probability that at least one of several receivers in a ground station network can sense a spacecraft downlink.

In addition to the visibility statistics, the database contains queries that produce weather related graphs at the AVM sites. For example, Figure 7 shows the minimum, maximum and average temperature (in Celsius) for each day of November 1996 at Mt. Lemmon. As another example, Figure 8 shows the probability distribution function (PDF) of wind-speeds at Mt. Lemmon for the same period of time. These plots are obtained by feeding the results of a query to a Microsoft Chart function. In the future, we plan to correlate atmospheric transmission to meteorological measurements to build a simple model for predicting atmospheric visibility at other sites.

5. CONCLUSIONS

A Microsoft Access database with a Visual Basic user interface was developed to store and process the large amount of atmospheric visibility data collected by the AVM project. The database allows us to extract useful information with just a few clicks of the mouse. Complex tasks such as joint CDFs from multiple sites can be obtained using

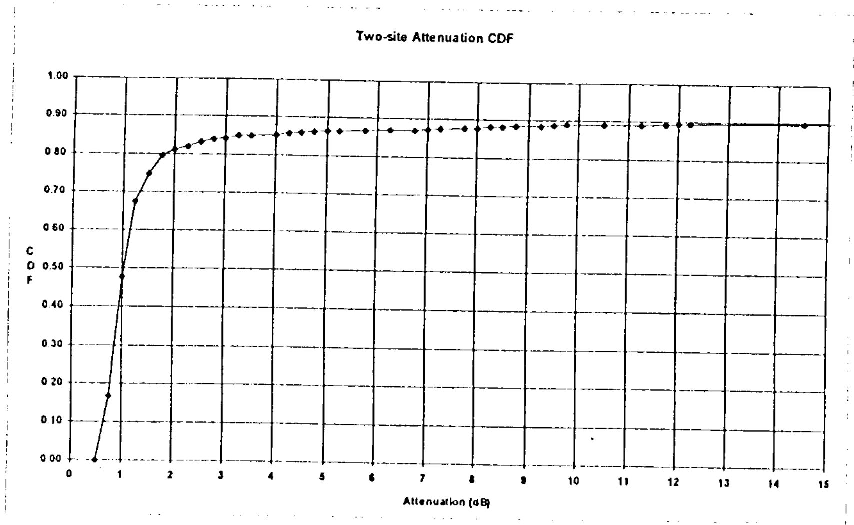

Figure 6. Two-site joint CDF for Mt. Lemmon and TMF during November 1996 at 532 nm.

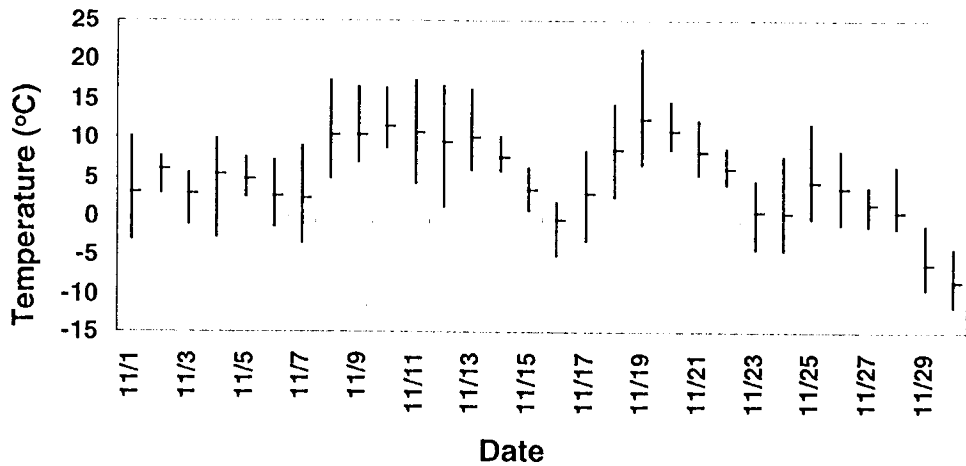

Figure 7. Plot showing the minimum, maximum and average temperature for each day in November 1996 at Mt. Lemmon.

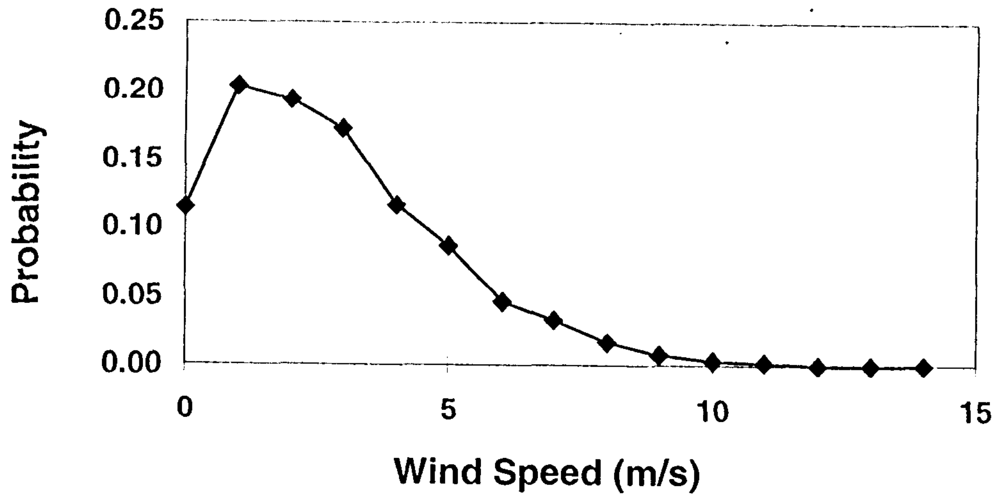

Figure 8. Probability distribution function of wind-speed at Mt. Lemmon in November 1996.

simple queries. In the future, we plan to include other weather related data (such as those from satellite observations) in the database. We, moreover, plan to develop statistical models for predicting atmospheric visibility at other sites through the use of existing meteorological data.

We have collected over three years of atmospheric visibility data from TMF and Mt. Lemmon. The AVM unit at Goldstone, on the other hand, has been operational little over a year and that too with severe communications problems. The planned upgrade of the CCD camera, computer and communication equipment (scheduled for completion in mid 1998) should enable gathering of higher quality data more reliably. The upgraded CCD camera in combination with a judicious choice of star list should enable us to obtain visibility data near the 1 micrometer wavelength and perhaps better daytime data as well.

ACKNOWLEDGMENTS

The work described in this article was performed by the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. The authors wish to thank all the personnel who have contributed to the project over the last several years especially D. Erickson, J. Lesh, T.-Y. Yan and G. Xu.

REFERENCES

1. N. Golshan, ed., *Proceeding of the Twentieth NASA Propagation Experiments Meeting (NAPEX XX)*, 1996. JPL Publication 96-20.
2. T.-Y. Yan, "Atmospheric transmission model through autonomous visibility measurement," *SPIE Proceeding* **2123**, pp. 400–8, 1994.
3. M. Jeganathan and D. E. Erickson, "Overview of the atmospheric visibility monitoring (AVM) program," *SPIE Proceeding* **2990**, pp. 114–20, 1997.
4. C. Sterken and J. Manfroid, *Astronomical Photometry: A Guide*, Kluwer Academic Publishers, 1992.
5. W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C*, Cambridge University Press, 1989.
6. A. Berk, L. Bernstein, G. Anderson, P. Acharya, D. Robertson, J. Chetwynd, and S. Adler-Golden, "MODTRAN cloud and multiple scattering upgrades with applications to AVIRIS," *Remote Sensing of Environment* , 1998. Accepted for publication.